AFRL-IF-RS-TR-2005-178
Final Technical Report
May 2005

ONTO-AGENTS-ENABLING INTELLIGENT AGENTS ON THE WEB

Stanford University

Sponsored by
Defense Advanced Research Projects Agency
DARPA Order No. K522

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

STINFO FINAL REPORT

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2005-178 has been reviewed and is approved for publication

APPROVED: /s/

NANCY A. KOZIARZ
Project Engineer

FOR THE DIRECTOR: /s/

JAMES W. CUSACK, Chief
Information Systems Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	MAY 2005	Final Jun 00 – Dec 04

4. TITLE AND SUBTITLE
ONTO-AGENTS-ENABLING INTELLIGENT AGENTS ON THE WEB

5. FUNDING NUMBERS
C - F30602-00-2-0594
PE - 62301E
PR - DAML
TA - 00
WU - 08

6. AUTHOR(S)
Gio Wiederhold, Rudi Studer,
Mark Musen, Stefan Decker and
Steffen Staab

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Stanford University
Computer Science Department
Gates Building
Stnaford California 94305

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington Virginia 22203-1714

AFRL/IFSA
525 Brooks Road
Rome New York 13441-4505

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2005-178

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Nancy A. Koziarz/IFSA/(315) 330-2828/Nancy.Koziarz@rl.af.mil

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.	

13. ABSTRACT ***(Maximum 200 Words)***
The objective of OntoAgents project was to develop concepts and modules that can serve as an ontology-driven 'Food Chain' for Advanced Applications on the Web. A number of tools were developed and made available to the community, primarily through webpages. Issues relating to acceptance, growth, and scalability of the semantic web appear in the conclusion.

14. SUBJECT TERMS
Ontology, Semantic Web, OWL, RDF, OnTo

15. NUMBER OF PAGES
23

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Table of Contents

List of Figures

Summary

The OntoAgents project was part of the DARPA-sponsored DAML effort (BAA 00-07). Our OntoAgents project started 1 July 2000 and terminated as of 31 Dec 2004. It was monitored by the Air Force Rome Laboratories (AFRL/IFSA, 525 Brooks Road, Rome, NY 13441-4505), the Air Force account is F30602-00-2-0594.
The cognizant Rome Laboratory staff are Nancy Koziarz (Nancy.Koziarz@rl.af.mil) and Mark J. Gorniak. DARPA program management included Jim Hendler and Murray Burke.

The Objective of OntoAgents project was to develop concepts and modules that can serve as an ontology-driven `Food Chain' for Advanced Applications on the Web.

Personnel

The principal investigator at Stanford was Prof. Gio Wiederhold and the principal scientific assistant project manager was Stefan Decker.

Gio Wiederhold retired formally in July 2001, but maintained responsibility for academic achievements, while recalled to 25% active duty for teaching and research. In July 2002 Stefan Decker and the principal focus of the project moved to the Information Science Institute (ISI) of the Univ. of Southern California (USC).

Until the summer of 2002 Stanford had a subcontract with Karlsruhe (Prof. Rudi Studer). After that date we had a subcontract for ongoing work with USC ISI (Stefan Decker) (ending 31 March 2004). USC also took over the Karlsruhe contract at that time. The Stanford extension beyond 1 April 2004 was to allow a student to complete his PhD thesis on the Ontology algebra.

We list the people that participated below, with their academic achievements during the project (in parenthesis) and their current positions. An asterisk (*) indicates that they received financial support from the OntoAgents project.

1. Prof. Gio Wiederhold, PhD, Principal Investigator * (Retired) Recalled for active duty to teach the Freshman course: Business on the Internet; consulting with MITRE Corp.
2. Prof. Rudi Studer, PhD, Principal Investigator, subcontract, Professor, University of Karlsruhe.
3. Mark Musen, Phd, MD, Co-Investigator (2001) Professor, Director Section of Medical Informatics, Stanford University
4. Stefan Decker, Project Leader * (PhD, Karlsruhe, January 2002) Research Staff, Digital Enterprise Research Institute (DERI) Galway, European Semantic Web Research Center and Nat.Univ. of Ireland.

5. Steffen Staab, PhD, Project Leader *(Habilitation, Karlsruhe, 2000), Professor, AI Institute (AIFB), University of Koblenz-Landau, Germany,
6. Sasha Buvac, research assistant (PhD 2004) Australian National University
7. Ray Fergerson, Information Systems Specialist * (2001-2002) Stanford Medical Informatics, Protégé project.
8. Siegfried Handschuh *(PhD, Karlsruhe, February 2005) AI Institute (AIFB), University of Karlsruhe, Germany,
9. Yuhui Jin *(MS, honors, Stanford 2003) Technical staff, Amazon.com, Seattle WA.
10. Maarten Kersten, PhD (2000, Visiting Researcher), CWI, Amsterdam, Holland
11. Martin Lacher, graduate researcher *tuition only (2001, Visiting Researcher) Technical Univ., Munich, Germany
12. Sergey Melnick, graduate researcher * (PhD, Leipzig, 2003) Microsoft Research, Redmond WA.
13. Prasenjit Mitra * (2001, PhD, Stanford, 2004) Assistant Professor, Penn State Univ., State College PA.
14. Natalya Fridman Noy, PHD (2003-2004) Senior Research Associate, Stanford Medical Informatics, Protégé project.
15. Sichun Xu, graduate research assistant *(2001, CS MS 2002) Ebay Corporation, CA.
16. Fernando Arguello, assistant (BS 2002, Santa Clara Univ.; participant in the Stanford SURF outreach program) Now at IBM Poughkeeps zXML group, NY.

Introduction

A notable aspect of the OntoAgents project is the broad interaction that it enabled among European and American researchers. As such it brought together extant and continuing research on the formal approaches to knowledge management, the pragmatic background of Expert systems approaches, and the concerns for scalability from database technologies.

Having a formal underpinning in complex projects is essential for reliability, maintainability to enable a long life, and scalability. Dealing with pragmatic issues is essential in dealing with practical situations, as heterogeneous data, autonomous participants, and effective performance. One example of attempting to bridge the gap is the proposal for Description logic programs: combining logic programs with description logic (DL) [GrosofHVD:03]. However, that combination only addresses the lowest level of DL proposed in the DAML setting. Another aspect is the concept and demonstration of an Ontology algebra. Such an algebra permits the interoperation of multiple, independently developed ontologies to interoperate in focused applications. When source ontologies change (as they will), the application ontology can be rapidly adapted using the existing algebraic formulation.

We do not claim that we solved these issues with finality. The tension between formality and scruffiness has been an issue in Artificial Intelligence since its inception, and will continue to hinder progress. The complexity of semantics is without bound, and progress will only uncover new depth that warrant research. We can only claim to have tried to make the semantic web community aware of the issues and provided constructive and well-founded directions.

Our vision was published as "An Information Food Chain for Advanced Applications on the WWW" [DeckerJMSW:00]. The diagram copied below depicts the approach and the different project parts. We will follow the process in our exposition.

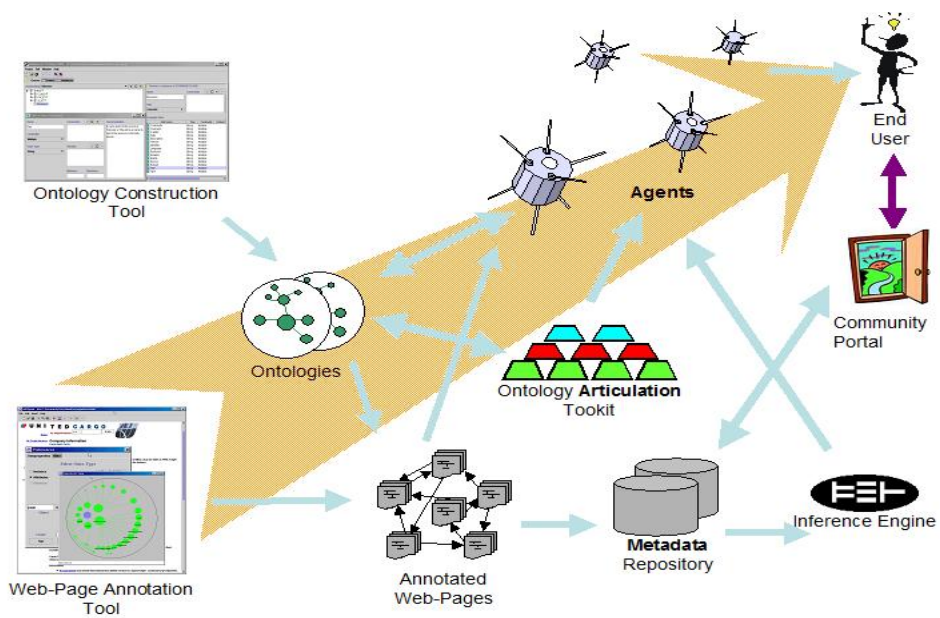

Figure 1 The semantic Web Foodchain

Methods, Assumptions, and Procedures

Annotation

To locate relevant pages on the semantic web they have to be annotated. Documents containing semantic annotations enable a more precise semantic search and allow for interoperation. These benefits, however, come at the cost of an increased authoring effort. In our work we have, therefore, presented a comprehensive framework which support users in dealing with the documents, the ontologies and the annotations that link documents to ontologies.

Manual annotation is tedious, and often done poorly. Even within the funded DAML project fewer pages were annotated than was hoped. In eCommerce, there has to be a sufficient business motivation to perform annotations, in then scientific world the motivation is less; although having the right tools will help [NoySDCFM:01]. Given the problems with syntax, semantics and pragmatics with annotation we identified the requirements of: consistency, proper reference, avoidance of redundancy, relational metadata maintenance, ease of use and efficiency [HandschuhSM:01] [CimianoHS:04].

Our work focused on methods to automate the annotation process [HandschuhS:03], using existing sources, as ontological knowledge [SureS:02], relational metadata [HandschuhS:02], [HandschuhSb:03] , digital libraries [MelnikGP:00], and other legacy data [VolzHSS:04]. We provide a comprehensive and pioneering annotation framework that reduces the complexity of Semantic Annotation for the annotator. The framework employs a comprehensive set of modules including inference services, crawler, document management system, ontology guidance/fact browser, and document editors/viewers. Process issues pertaining to the annotation/authoring task are modularized from content descriptions by a meta ontology.

The framework has been *prototypically implemented* in the open source project OntoMat hosted by the DARPA DAML program [http://projects.semwebcentral.org/projects/ontomat/]. The annotation framework is populated with specialized methods for:

- Manual Annotation: The transformation of existing document resources, into relatable knowledge structures which represent the underlying information.
- Authoring of Documents: Authoring lets users create metadata with little added effort, while putting together the content of a page.
- Semi-automatic Annotation: Semi-automatic Annotation based on Information Extraction.
- Deep Annotation: Considers Web pages which are generated from a database by annotation of the underlying database.

The size of the deep web has been estimated to be many times larger than the shallow web, the directly accessible information as retrieved by tools as Google. The deep web covers the information dynamically populated from databases, as typically done by business services, and such important to the future of the semantic web [HandschuhSa:03] [HandschuhSV:03a] [HandschuhSV:03b] [HandschuhSV:03]. Its effective size is hard to measure, since the same database -- say stock prices -- van be provided by multiple services. Measurements of the deep web have also counted the huge volume of images that satellites have collected. While those are also accessible on the web, the value in terms of actionable information per megabyte stored is small. But no matter what the size metric should be, dealing with deep web will be crucial, and require tools that are linked to database technology.

OntoMat is the reference implementation of the CREAM framework [HandschuhSM:01] [HandschuhSC:02]. It is Java-based and provides a plug-in interface for extensions for further applications. It has been used in several cases, e.g. the annotation of paper abstracts for the International Conferences on Semantic Web (ISWC 2002, 2003, 2004) by each of the authors. Ontomat is in use on class room machines in an obligatory

Semantic Web course for informatics students in Prague, which enrolls some 250 people every year [http://nb.vse.cz/~svatek/modz.htm].

Ontologies

Information for annotation can be derived from many sources, as discussed above, but require tools to create effective ontologies [MaedcheS:03] [MaedcheNS:03] [OberleVSM:04] [StaabEAD:01] . Automation, using AI learning technologies is one approach [MaedcheS:01] [MaedcheS:03]. Ontological information may be obtained by inferencing [SureAS:03] [SureEASSW:02].
Once ontologies are established they have to be maintained [AbererEa:04]. The ontologies can be stored anywhere on a dynamic network [NejdlEa:02], or on a grid [TangmunarunkitDK:03].

The core Protege system software was modified to support the development of RDF enhancements to Protégé. In order to allow ontologies maintained within the Protégé system to interoperate with the RDF representation, a plugin is available from the Protégé web site [NoySDCFM:01].

Much of this information is summarized in a handbook, to which most OntoAgents project participants have contributed [StaabS:04]. A future research challenge is developing support systems for ontology evolution and supporting adaptation of the applications that use those ontologies, when the ontologies are updated [MitraWD:01] [Oliver:00].

Knowledge Management

Having well structured and focused ontologies provides a basis for organizing knowledge, the main distinguishing property in modern organizations and businesses [StaabSS:02] [StaabSSS:01] [SureS:02] [SureSS:02]. A complementary current approach to knowledge management are topic boards, and we explored that relationship [LacherD:01]. Organizing knowledge encapsulated in governmental regulations is a current issue as well, here we are cooperating with a project in Stanford's department of Civil and Environmental Engineering [LauKLW:04] [LauLW:05].

The commonality that can be achieved is still unclear, but a topic of continuing research [BernsteinHJRW:00] [Melnik:03] [MelnikRBa:03] [MelnikRBb:03].

Infrastructure

Web services are expected to operate in a widely distributed environment, and we interacted with and supported projects that focused on the required infrastructure [LiuPLW:04] [MelnikD:00] [MelnikGR:02]. The scalability of these systems, while maintaining correctness is a major concern, as expressed in a workshop that was organized by ISI colleagues [VolzDC:03]. We investigated how Semantic Web standards such as RDF and OWL can be used within our reasoning language TRIPLE [SintekD:03] for resource matching for the Grid. Results are promising and have spawned follow up

work in the resource matchmaking area using rules. The second topic was centered around the emerging Grid notion: in [TangmunarunkitDK:03] and [HarthDHTK:04].

Furthermore we were involved in discussions around a potential rules standard for the Semantic Web [HorrocksADKGW:03]. Extended operators, such as aggregation in TRIPLE, proved to be necessary for matchmaking applications. TRIPLE is under continued development and is available at http://triple.semanticweb.org.

Inferencing Agents

Application of the knowledge, through agents that perform reasoning and inferencing procedures, is central to the promise of the semantic web. As implied in the introduction to this section, it is here where the technologies now used by the AI community need to come together. Scalability and pragmatic effectiveness are expected in the semantic web.

Inferencing, i.e., relating the knowledge sentences from the sources to achieve higher level goals, is needed during construction of onotogies [SureAS:02] as well as during their application [NoySDCFM:01]. Work within OntoAgents has focused on TRIPLE [DeckerS:02] [SintekD:03], which shares the RDF and OWL-DL knowledge representations with other DAML projects. TRIPLE's Horn-logic-based approaches have been applied to representations used for description logics [GrosofHVD:03]. That work identified the common intersection between Logic Programming languages and Description Logic languages, and dubbed it Description Logic Programming. We showed that a large part of a language such as OWL or DAML+OIL can be captured within that Description Logic Programming framework, which allows for efficient reasoners for these language subsets. That work is now widely cited and used in follow up work.

An underlying issue is how demanding the applications of the semantic web will be. If use is no more complex than seen in the common search models today, available technologies will provide adequately broad information, but not avoid the dreaded information overload. Any excess or wrong information must now be filtered out by the end-users. Annotation will improve that filtering somewhat [AgarwalHS:03]. But semantic web proponents expect a much greater level of automation. For routine business applications filtering has to be carried without user participation. More complex, multi-service applications require a greater depth of inferencing to obtain adequate information; but filtering of mismatches is essential to avoid overload. The optimal, or at least effective tradeoff between missing some information and receiving excessive junk must be based on a situational criterion, that balance warrants formal quantification [Wiederhold:02].

Added value for applications is generated when knowledge can be applied to projects outside of the computer science community. A major test of today's capabilities was the Halo Project. [FriedlandEa:04] [FriedlandEa:05]. Participants from Karlsruhe, using simple deductive inferencing, were able to compete effectively using fewer resources and less time. The approach used by their Ontoprise system required far less tuning and narrowing of the knowledge bases than approaches used by others [AngeleMOSW:03].

Testing Resources

For testing purposes, we have made a large and densely linked XML file of Movies, their directors, actors, casts, and remakes (for deeper inferencing) available [WiederholdA:2004]. This material could be converted to RDF, and provide a more complex setting than the bibliographic files now in common use. A tool, XLint, was developed to syntactic report errors in XML files to allow bulk repairs of systematic errors to proceed rapidly [ArguelloCW: 04]. Systematic errors will occur when converting large HTML data collections to XML, because of the strict well-formedness constraints imposed by XML.

The OntoAgents project also supported a RDF-encoding of Wordnet 1.6 [MelnikD:01]. This RDF resources is an input to the W3C Best Practices Working Group.

Resolving Heterogeneity

An issue of concern is that as the web grows, many ontologies will evolve, exacerbating issues of scalability [BozsakEa:02] [VolzDC:03]. When applications require information from multiple autonomous sources, we cannot expect a common ontology, since a joint or global ontology would hinder growth and effectiveness in narrow domains [Wiederhold:03]. The differences may be minor, but their import is hard to assess by users, unless tools are made available [MaedcheS:02].

Resolving semantic heterogeneity among knowledge and data resources has been an issue of research at Stanford for some time [Wiederhold:94] [Wiederhold:00] [MitraWK:00] [Melnik:00]. Our concepts focus on an Ontology algebra. Our focus within the OntoAgents project has been on the articulation of pairs of ontologies using semi-automated methods [MitraWK:00] [MitraWD:01] [Wiederhold:01] [MitraW:02] [Mitra:04] [MitraW:04].

Each articulation can focus on a specific application, and becomes easier to maintain and manage. An initial phase suggests articulation rules, containing candidate matches for interoperation. When validated by the interoperation expert they enter an application-focused repository. During the operational phase interoperation among resources described by those articulation rules can proceed automatically.

Some related work at Stanford is quite formal, but has provided important background [McCarthy:93] [Buvac:04].

Portals

Access to knowledge and information is provided through portals, the desktop interfaces used by the public to interact with the web [DeckerF:04]. Consistency and maintainability demands that those portals are driven by ontologies [MaedcheSSS:01] [JinDW:01] [JinX:02] [StaabSSV:02] [MaedcheSSSS:03] [MaedcheSSSV:02] [HartmannS:04]. While promising, the Stanford effort on ontology-based assistance in the construction of portals was only brought to a prototype stage.

Web services

Obtaining actionable information from the web services is the end objective envisaged in OntoAgents, as well as inn the entire DAML effort [MaedcheNS:03]

The business model of web services is just now being established. It is unclear how these services will be supported in the long term, by the sale of associated products, by advertising, by volunteer efforts, or by public funds, but it will likely be a combination of all of these [AgarwalHS:03]. When the product of the web service is information, as now kept in databases, subscription models are common, but reduce flexibility. Interaction with the database [AngeleMOSW:03] [DeckerK:03] and digital library [MelnikGP:00] [LarsenEa:03] [Wiederhold:03a] [Wiederhold:03b] communities is important for management of content.

The lack of experience with semantic web operations makes it difficult to formalize a business model, even though business-oriented metrics will be essential to gain support [Wiederhold:05].

Results and Discussion

We cite here the web sites that contain results from the Ontoagents Peoject. The References cite a large number of publications where the OntoAgents project provided some input or relevance. The modest investment in the ongoing work at the University of Karlsruhe was especially productive. Not all of the papers listed in the references are cited in the descriptive text above. A number of workshops were organized as well.

Websites

Information about OntoAgents, its products, and related research is available at

http://www-db.stanford.edu/OntoAgents/ = OntoAgents abstracts only [Decker]

http://www.semanticweb.org/ = General web site, not updated since June 2003 [Decker et al.]

http://annotation.semanticweb.org/ = Web site dedicated to semantic annotation [Handschuh]

http://projects.semwebcentral.org/projects/ontomat/ = Project page and cvs repository of Ontomat OWL/RDF semantic annotation tool , available under the GNU Lesser General Public License (LGPL) [Handschuh]

http://projects.semwebcentral.org/projects/owlcrawler/ = Project page and cvs repository of OWL/RDF or FOAF crawler [Handschuh]

http://www.aifb.uni-karlsruhe.de/about.html = The SSEAL portal at the AIFB Karlsruhe.

http://www-db.stanford.edu/SKC/index.html = Predecessor project on Semantic Interoperation [Mitra, Wiederhold]

http://www.aifb.uni-karlsruhe.de/WBS/sha = Ontology development [Handschuh]

http://protege.stanford.edu/plugins/rdf/ = Protégé RDF backend plugin [Fergerson]

http://www-db.stanford.edu/OntoAgents/xlint/index.html = Xlint processor [Arguello]

http://www.dfki.uni-kl.de/frodo/triple and http://triple.semanticweb.org = TRIPLE inference engine [Decker and Sintek]

http://www.ontoweb.org/download/deliverables/D21_Final-final.pdf = Scenarios [Leger et al.]

http://edutella.jxta.org/ = RDF-based Metadata Infrastructure for P2P Applications (PADLR/Edutella)

Conclusions

This Section represents my personal observation on three topics, and reflects in no way the work and opinions of other DAML or OntoAgent project participants. I have received some valuable feedback from OntoAgents researchers. Since my participation diminished greatly after my retirement I will not be aware of all advances made since then. So, if issues I list below have been overcome, congratulations!

The DAML project was initiated at the birth of the semantic web. It contributed greatly to define a new research area, but, because of its novelty, also had to depend on researchers that had been active earlier in other computer science settings. As a result, some tradeoffs to bring the semantic web, as envisaged here into practical real-world use, have not been established as well as the need to be.

Robustness.

Acceptance of RDF or similar representations is today a major barrier for users outside of academia, who today are still fighting XML and its requirements. In reviewing web technology we observe a trend.

The acceptance of HTML was enabled by the robustness of the browsers. Even today many HTML page on the web have syntactic and content errors, but they remain human-understandable, and can also be adequately processed by search engines screenscrapers. However, a single syntactic error in an XML document typically prevents access to all subsequent information. Such a punitive interpretation is discouraging. RDF seems to be no better. It is unclear to what extent the problem can be addressed by improving the representation versus adapting the interpreters. Some settings of the semantic web indeed require completeness and the attendant cost to attain perfection; but many do not. When searching a hotel I am happy with a dozen choices, any more creates overload. it is unlikely that the 13th hotel choice, not shown properly because of a syntax error, would significantly change my decision. If that hotel entry had been early in an XML list, however, I would have failed to see all of the remainder. Can the expected perfection become a parameter?

Automatic annotation

Annotation is crucial to the concept of the semantic web, but also time-consuming. There has been much research here, but I have not yet seen any public

business webpages that were annotated using such tools. Without applications that allow the providers to profit from the annotations, there is little benefit and actually some risk of misuse of annotations. Webpages used to improve internal knowledge management can, and are profitably annotated in some organizations.

For legacy web pages automatic assistance for annotation is essential and must be convenient rather than perfect. The first round provided by automation should be easy, maybe even invisible to the users. Its output should allow convenient refinement, by humans as well as tools. That will likely require tracking of the provenance of annotations, so we don't repeat the validation problems now encountered in the genome project.

New technologies are emerging that provide annotations as the data are entered. Interoperation of those annotations will require that those technologies use the same ontologies; or that the ontologies themselves become interoperable. There are justifiable barriers to sharing ontologies at the level of the creators of the data, that will not be overcome by presenting a vision of a grand future [Wiederhold:02] [Wiederhold:03a]. If there are inadequate benefits compared to the costs for the information generator, then the imposition of external expertise, supported by the users that benefit, has to be enabled.

One problem is that an optimal ontology for one application category, as geo-coding for photographic images (FOAF), is not likely to be effective for geo-coding of Marine Corps logistic destinations and interchange points [Berg:03].

Any annotation must be viewable, else no feedback will be generated by owners and users. If annotations remain disjoint, (obsolete) computer-science principles may be served, but failures due to annotation errors will remain mysteries. The lack of integration of annotation and viewable content is a major discouragement in current implementations.

Recommendation

For dissemination of DAML and successor results, the potential customers of those results need to see the effectiveness of research products in an easy-to-perceive and relatively unbiased manner. Having some publicly available, realistic and compelling scenarios will also focus semantic web research, since they can be used by the community to test their work, This suggestions is not original, and was widely discussed in 2002, when it was obvious that using the DAML machinery merely to conclude that "Mary is the parent of Bill" was not compelling [Pease:02] [Brachman:02].

There was a nice scenario in the Berners-Lee, Hendler and Lassila Scientific American Article, but I have not seen it actually demonstrated. That scenario is quite ambitious, and depends too much on resources that do not exist today. Other example scenarios have been listed on DAML participant reports, but not worked out, as far as I know, to provide a sharable set of test cases. The European OntoWeb Project lists 21 `Successful Scenarios' of Semantic Web technology, but none is documented yet to the level that it can be used as a test case for measuring semantic web technology progress and innovation.

The relevant site data also have to be available. The Halo project provided that basis, in the area of answering questions on High-school level Chemistry. Its creation comprised much of the cost of the Halo project. The DARPA community did use scenarios in the prior HPKB project and provided data for participants in its TREC efforts. The Database community now has its standard transaction streams used to assess progress.

Having standard scenarios, of varying types, with substantial data ,will allow the community to assess open issues, as the tradeoff among formality and scruffiness needed in semantic web engines, and the failure rates and performance issues faced by alternate logics.

References

[AbererEa:04] Karl Aberer, Philippe Cudré-Mauroux, A. Ouksel, T. Catarci, M.-S. Hacid, A. Illarramendi, V. Kashyap, M. Mecella, E. Mena, E. Neuhold, O. De Troyer, T. Risse, M. Scannapieco, F. Saltor, L. De Santis, S. Spaccapietra, Steffen Staab, Rudi Studer: "Emergent Semantics Principles and Issues (invited contribution)"; Y. Lee et al. (Eds.) *Proceedings of Database Systems for Advances Applications, 9th International Conference, DASFAA 2004*, Jeju Island, Korea, March 17-19, 2004, LNCS 2973 Springer. pp. 25-38.

[AgarwalHS:03] S. Agarwal, Siegfried Handschuh, Steffen Staab: "Surfing the Service Web"; D. Fensel and K. Sycara and J. Mylopoulos: *Proc. of ISWC-2003* - International Semantic Web Conference, LNCS 2870, Springer 2003.

[AngeleMOSW:03] J. Angele, E. Moench, H. Oppermann, Steffen Staab, and D. Wenke: "Ontology-Based Query and Answering in Chemistry: OntoNova @ Project Halo. Proc. of ISWC-2003 - International Semantic Web Conference, Sanibel Island, Florida, October 2003, LNCS, Springer 2003.

[ArguelloCW: 04] Fernando Arguello, Vincent Chu, and Gio Wiederhold: XLint: a XML preprocessor; http://www-db.stanford.edu/OntoAgents/xlint/index.html .

[Berg:03] Murray Berg (ed): Ontology Driven Knowledge Dissemination to Support Marine Corps Logistics; Hot DAML, newsletter 10, April 2003.

[BernsteinHJRW:00] Philip A. Bernstein (moderator), Laura Hass, Matthias Jarke, Erhard Rahm, Gio Wiederhold (panelists): Is Generic Metadata Management Feasible? Panel, VLDB 2000

[BozsakEa:02] A. Bozsak, M. Ehrig, Siegfried Handschuh, A. Hotho, A. Mädche, B. Motik, D. Oberle, C. Schmitz, Steffen Staab, L. Stojanovic, N. Stojanovic, Rudi Studer, G. Stumme, Y. Sure, J. Tane, R. Volz, V. Zacharias: "KAON - Towards a large scale Semantic Web"; *Proceedings of EC-Web 2002*. Aix-en-Provence, France, September 2-6, 2002. LNCS, Springer, 2002.

[Brachman:02] Ron Brachman: Request for input to motivate research on Cognitive systems; DARPA, August 2002.

[Buvac:04] Sasha Buvac: Modality in the Ways of Contexts; Stanford University PhD thesis, 2004.

[CimianoS:03] P. Cimiano, S. Handschuh: "Ontology-based Linguistic Annotation"; *Proc. of the ACL Workshop on Linguistic Annotation*, Sapporo, Japan, 2003.

[CimianoHS:04] P. Cimiano, Siegfried Handschuh, Steffen Staab: "Towards the Self-Annotating Web. In *Proceedings of the 13th International World Wide Web Conference, WWW 2004*, New York, USA, May, 2004. ACM Press.

[CruzDEM:02] Isabel F. Cruz, Stefan Decker, Jirtme Euzenat, and Deborah McGuinness (eds.): *The Emerging Semantic Web;* Proceedings of SWWS'02, Vol. 75, Frontiers in Artificial Intelligence and Applications, IOS Press, Amsterdam, NL June 2002.

[CruzKDE:03] Isabel F. Cruz, Vipul Kashyap, Stefan Decker, Rainer Eckstein (eds.): *Proceedings of SWDB'03, The first International Workshop on Semantic Web and Database;* Co-located with VLDB 2003, Humboldt-Universitaet, Berlin, Germany, September 7-8, 2003 SWDB 2003.

[DeckerF:04] Stefan Decker, Martin R. Frank: "The Networked Semantic Desktop"; WWW *Workshop on Application Design, Development and Implementation Issues in the Semantic Web* 2004.

[DeckerGHV:03] Stefan Decker, B. N. Grosof, I. Horrocks, R. Volz, "Description Logic Programs: Combining Logic Programs with Description Logic," in *Proc. 12th International World Wide Web Conference*, 2003,

[DeckerJMSW:00] Stefan Decker, Jan Janninck, Prasenjit Mitra, Rudi Studer, and Gio Wiederhold: "An Information Food Chain for Advanced Applications on the WWW"; *ECDL 2000*, Proc. 4th European Conference on Research and Advanced Technology for Digital Libraries, Springer LCNS Vol.1923, 2000, pages 490-497; http://www-db.stanford.edu/pub/gio/2000/FoodChain.doc

[DeckerK:03] Stefan Decker, Vipul Kashyap: "The Semantic Web: Semantics for Data on the Web"; *VLDB 2003*:tutorial, p. 1148.

[DeckerMVFKBEH:00] Stefan Decker, Sergey Melnik, F. Van Harmelen, D. Fensel, M. Klein, J. Broekstra, M. Erdmann, and I. Horrocks: "The Semantic Web: The roles of XML and RDF. *IEEE Internet Computing*, 2000, October, *15 (3)* , pp. 63-74; http://www.cs.vu.nl/~frankh/postscript/IEEE-IC00.pdf

[DeckerS:02] Stefan Decker and M. Sintek: "Triple—a query, inference, and transformation language for the semantic web"; 1st International Semantic Web Conference (ISWC), June 2002, http://triple.semanticweb.org

[JinDW:01] Yuhui Jin, Stefan Decker, Gio Wiederhold: "OntoWebber: Model-Driven Ontology-Based Web Site Management"; 1st International Semantic Web Working Symposium (SWWS `01), Stanford University, Stanford, CA, July 29-Aug 1, 2001;http://www-db.stanford.edu/pub/gio/2001/Ontowebber01.pdf.

[JinX:02] Yuhui Jin, Suchun Xu: "Managing Web Sites with OntoWebber", unpublished MS, Stanford OntoAgents project, 2002; abstract at http://www-db.stanford.edu/OntoAgents/

[LiuPLW:04] David Liu, Jun Peng, Kincho H. Law, and Gio Wiederhold: Efficient Integration of Web Services with Distributed Data Flow and Active Mediation; <I>ICEC'04 conference</I>, Delft, The Netherlands, Oct.2004; <A HREF="http://www-db.stanford.edu/pub/gio/2004/ICEC04reduced.doc">

[FriedlandEa:04] Noah S. Friedland, Paul G. Allen, Michael Witbrock, Gavin Matthews, Nancy Salay, Pierluigi Miraglia, Jurgen Angele, Steffen Staab, David Israel, Vinay Chaudhri, Bruce Porter, Ken Barker, Peter Clark: "Towards a Quantitative, Platform-Independent Analysis of Knowledge Systems"; *Proceedings of the Conference on Knowledge Representation and Reasoning - KR-2004*. AAAI Press, 2004. **Nominated for Best Paper Award.**

[FriedlandEa:05] N. Friedland, P. Allen, G. Matthews, M. Witbrock, D. Baxter, J. Curtis, B. Shepard, Prasenjit Miraglia, J. Angele, Steffen Staab, E. Moench, H. Oppermann, D. Wenke, D. Israel, V. Chaudhri, B. Porter, K. Barker, J. Fan, S. Chaw, P. Yeh, D. Tecuci, P. Clark: "Project Halo: Towards a Digital Aristotle"; *AI Magazine*, 2005.

[GrosofHVD:03] Benjamin N. Grosof, Ian Horrocks, Raphael Volz, Stefan Decker: "Description logic programs: combining logic programs with description logic"; *WWW 2003*: 48-57

[HandschuhS:02] Siegfried Handschuh, Steffen Staab: "Authoring and Annotation of Web Pages in CREAM"; *Proceedings of the 11th International World Wide Web Conference, WWW 2002*, Honolulu, Hawaii, May 7-11, 2002. ACM Press; http://www2002.org/CDROM/refereed/506/index.html

[HandschuhS:03] Siegfried Handschuh, Steffen Staab (eds.): Annotation for the Semantic Web; IOS Press, 2003.

[HandschuhS:03a] Siegfried Handschuh, Steffen Staab: "Annotation of the Shallow and the Deep Web"; Siegfried Handschuh & Steffen Staab (eds.), *Annotation for the Semantic Web.* IOS Press, 2003.

[HandschuhS:03b] Siegfried Handschuh, Steffen Staab: "CREAM - CREAting Metadata for the Semantic Web"; *Computer Networks*. Vol.42. pp.579-598, Elsevier 2003.

[HandschuhSC:02] Siegfried Handschuh, Steffen Staab, F. Ciravegna: "S-CREAM - Semi-automatic CREAtion of Metadata"; *Proc. of the European Conference on Knowledge Acquisition and Management - EKAW-2002*. Madrid, Spain, October 1-4, 2002. LNCS, Springer, 2002; *Expert Update*, vol.5 No.3 Autumn 2002.

[HandschuhSM:01] Siegfried Handschuh, Steffen Staab, Alexander Maedche: "CREAM — Creating relational metadata with a component-based, ontology-driven annotation framework"; S*emantic Web Workshop*, July 2001; *ACM K-CAP 2001*. October, Vancouver. http://www.semanticweb.org/SWWS/program/full/paper4.pdf

[HandschuhSM:02] Siegfried Handschuh, Steffen Staab, Alexander Maedche: "CREAting relational Metata (CREAM) -- a framework for semantic annotation"; I. Curz, S. Decker, J. Euzenat (Eds.): *The Emerging Semantic Web*, IOS Press, 2002.

[HandschuhSS:03] Siegfried Handschuh, Steffen Staab, Rudi Studer: "Leveraging metadata creation for the Semantic Web with CREAM"; R. Kruse et al. (Eds.), KI '2003 - Advances in Artificial Intelligence. Proc. of the Annual German Conference on AI, Hamburg, September, 2003, LNAI. Berlin, Heidelberg: Springer.

[HandschuhSV:03a] Siegfried Handschuh, Steffen Staab, R. Volz: "On Deep Annotation"; *Proceedings of the 12th International World Wide Web Conference, WWW 2003*, Budapest, Hungary, May 20-24, 2003. ACM Press.

[HandschuhSVb:03b] Siegfried Handschuh, Steffen Staab, R. Volz: "Annotation for the Deep Web"; *IEEE Intelligent Systems*, Vol.18 No.5, Sep/Oct 2003, pp. 42-48 (Special issue on information integration).

[HarthDHTK:04] Andreas Harth, Stefan Decker, Yu He, Hongsuda Tangmunarunkit, Carl Kesselman: "A Semantic Matchmaker Service on the Grid"; *WWW (Alternate Track Papers & Posters) 2004*: 326-327.

[HartmannS:04] Jens Hartmann, York Sure: "An Infrastructure for Scalable, Reliable Semantic Portals"; IEEE Intelligent Systems 19 (3): 58-65. May 2004.

[HorrocksADKGW:03] Ian Horrocks, Juergen Angele, Stefan Decker, Michael Kifer, Benjamin N. Grosof, Gerd Wagner: "Where Are the Rules?"; *IEEE Intelligent Systems* 18(5): 76-83 (2003)

[LacherD:01] Martin S. Lacher and Stefan Decker: "On the Integration of Topic Maps and RDF Data"; first Semantic Web Workshop, Stanford July 2001; http://www.semanticweb.org/SWWS/program/full/paper53.pdf .

[LarsenEa:03] Ron Larsen, Howard Wactlar, et al.: Knowledge lost in Information; Report of the NSF Workshop on Digital Library Research Directions; June 15-17, 2003, Chatham, MA.

[LauKLW:04] Gloria T. Lau, Shawn Kerrigan, Kincho H. Law, and Gio Wiederhold: An E-Government Information Architecture for Regulation Analysis and Compliance Assistance; *ICEC'04 conference*, Delft, The Netherlands, Oct.2004.

[LauLW:05] Gloria T. Lau, Kincho H. Law, and Gio Wiederhold: A Relatedness Analysis Tool for Comparing Drafted Regulations and Associated Public Comments*; I/S: A Journal of Law and Policy for the Information Society*, OSU & CMU, Vol.1 No.1, 2005.

[LiuPLW:04] David Liu, Jun Peng, Kincho H. Law, and Gio Wiederhold: Efficient Integration of Web Services with Distributed Data Flow and Active Mediation; *ICEC'04 conference*, Delft, The Netherlands, Oct.2004;

[MaedcheSSS:01] A. Mädche, Steffen Staab, N. Stojanovic, Rudi Studer, Y. Sure: "SEAL - A Framework for Developing SEmantic portALs"; *BNCOD 2001 - 18th British National Conference on Databases*. Oxford, UK, 9th - 11th July 2001, LNCS 2097, Springer Verlag, 2001, pp. 1-22.

[MaedcheSSSS:03] Alexander Maedche, Steffen Staab, N. Stojanovic, Rudi Studer, Y. Sure: "SEmantic portAL - The SEAL approach"; *Spinning the Semantic Web*. D. Fensel, J. Hendler, H. Lieberman, W. Wahlster (eds.) MIT Press, Cambridge, MA, 2003.

[MaedcheNS:03] Alexander Mädche, G. Neumann, Steffen Staab: "Bootstrapping an Ontology-based Information Extraction System"; P. Szczepaniak, J. Segovia, J. Kacprzyk, L. Zadeh (eds.) *Intelligent Exploration of the Web*. Springer / Physica Verlag, Heidelberg, 2003.

[MaedcheS:01] Alexander Maedche and Steffen Staab, "Ontology Learning for the Semantic Web," *IEEE Intelligent Systems*, 16, 2, 2001.

[MaedcheS:02] Alexander Mädche, Steffen Staab: "Measuring Similarity between Ontologies"; *Proc. Of the European Conference on Knowledge Acquisition and Management - EKAW-2002*. Madrid, Spain, LNCS, Springer, 2002.

[MaedcheS:03] Alexander Maedche, Steffen Staab: "KAON: The Karlsruhe Ontology and Semantic Web Meta Project"; *Künstliche Intelligenz.* Special Issue on Semantic Web. 3/2003, pp. 27-30.

[MaedcheSSSV:02] Alexander Maedche, Steffen Staab, Rudi Studer, Y. Sure, R. Volz: "SEAL — Tying Up Information Integration and Web Site Management by Ontologies"; *IEEE Data Engineering Bulletin*, 2002.

[McCarthy:93] John McCarthy: "Notes on Formalizing context"; Proceedings of the 13th International Joint Conference on Artificial Intelligence, pages 555–562, 1993.

[Melnik:00] Sergey Melnik: "Declarative mediation in distributed systems"; *Proceedings of the International Conference on Conceptual Modeling (ER'00),* 2000.

[Melnik:03] Sergey Melnik: Generic Model Management; PhD thesis, Univ. of Leipzig, 2003.

[MelnikD:00] Sergey Melnik and Stefan Decker: "A Layered Approach to Information Modeling and Interoperability on the Web"; *Proc. ECDL'00 Workshop on the Semantic Web,* Lisbon, Portugal, 2000.

[MelnikGP:00] Sergey Melnik, Garcia-Molina, H., and Paepcke, A: "A Mediation Infrastructure for Digital Library Services"; *Proc. of the 5th ACM Intl. Conf. on Digital Libraries*. 2000. 123–132.

[MelnikGR:02] Sergey Melnik, Garcia-Molina, H., and Rahm, E.: "Similarity Flooding: A Versatile Graph Matching Algorithm and Its Application to Schema Matching"; *Proc. of 12th Intl. Conf. on Data Engineering (ICDE)*. IEEE Computer Society, February 2002, 117–128.

[MelnikRBa:03] Sergey Melnik, Rahm, E., and Bernstein, P. A. : "Developing Metadata-Intensive Applications with Rondo"; *Intl. Journal on Web Semantics*. 2003.

[MelnikRBb:03] Sergey Melnik, Rahm, E., and Bernstein, P. A. : "Rondo: A Programming Platform for Generic Model Management"; *Proc. of ACM SIGMOD Intl. Conf. on Management of Data, 2003*.

[Mitra:04] Prasenjit Mitra: An Algebraic Framework for the Interoperation of Ontologies; Stanford University PhD thesis, 2004, http://www-db.stanford.edu/~prasen9/thesis-pm.pdf .

[MitraW:02] Prasenjit Mitra, and G. Wiederhold: "An Algebra for the Composition of Ontologies"; Workshop on Knowledge Transformation for the Semantic Web. (2002). http://www.cs.vu.nl/~borys/events/ktsw2002.pdf

[MitraW:04] Prasenjit Mitra and Gio Wiederhold: "An Ontology-Composition Algebra"; S.Staab, R.Studer (eds.): *Handbook on Ontologies*, Springer Series: International Handbooks on Information Systems, 2004, pages 93-113.

[MitraWD:01] Prasenjit Mitra, Gio Wiederhold and Stefan Decker: "A Scalable Framework for Interoperation of Information Sources"; *1st International Semantic Web Working Symposium (SWWS `01),* Stanford University, Stanford, CA, Jul. 2001; http://www.semanticweb.org/SWWS/program/full/paper51.pdf.

[MitraWK:00] Prasenjit Mitra, Gio Wiederhold, and Kersten, M. L. : "A Graph-Oriented Model for Articulation of Ontology Interdependencies"; *Proc. of Intl. Conf. on Extending*

Database Technology (EDBT), C. Zaniolo, P. C. Lockemann, M. H. Scholl, and T. Grust, Eds. Lecture Notes in Computer Science, vol. 1777. Springer, 86–100. 2000.

[NejdlEa:02] Wolfgang Nejdl, Boris Wolf, Changtao Qu, Stefan Decker, Michael Sintek, Ambjörn Naeve, Mikael Nilsson, Matthias Palmér, Tore Risch: EDUTELLA: a P2P networking infrastructure based on RDF. *WWW 2002*: 604-615

[NoySDCFM:01] N. F. Noy, M. Sintek, Stefan Decker, M. Crubezy, R. W. Fergerson, and M. A. Musen, "Creating Semantic Web Contents with Protege-2000," *IEEE Intelligent Systems*, Vol.16 No.2, 2001, special issue on Semantic Web, pp. 60-71.

[OberleSSV:04] D. Oberle, Steffen Staab, Rudi Studer, R. Volz: "Supporting Application Development in the Semantic Web"; *ACM Transactions on Internet Technology*, 4(4), 2004.

[OberleVSM:04] D. Oberle, R. Volz, Steffen Staab, B. Motik: "An Extensible Ontology Software Environment"; Steffen Staab & Rudi Studer (eds.) *Handbook on Ontologies*. Springer 2004.

[Oliver:00] Diane E. Oliver: Change Management and Synchronization of Local and Shared Versions of a Controlled Vocabulary; PhD thesis, Stanford University, 2000.

[Pease:02] Adam Pease: Why Use DAML; DAML working paper, 10 April 2002.

[PekarS:03] V. Pekar, Steffen Staab: "Word classification based on combined measures of distributional and semantic similarity"; *Proc. Research Notes of the 10th Conference of the European Chapter of the Association for Computational Linguistics*, April 2003, Budapest, Hungary

[Schlosser et al, 2002a] Mario T. Schlosser, Michael Sintek, Stefan Decker, Wolfgang Nejdl: HyperCuP - Hypercubes, Ontologies, and Efficient Search on Peer-to-Peer Networks. *AP2PC 2002*: 112-124

[Schlosser et al, 2002b] Mario T. Schlosser, Michael Sintek, Stefan Decker, Wolfgang Nejdl: A Scalable and Ontology-Based P2P Infrastructure for Semantic Web Services. *Peer-to-Peer Computing 2002*: 104-111

[SintekD:03] Michael Sintek, Stefan Decker: "Using TRIPLE for Business Agents on the Semantic Web"; *Electronic Commerce Research and Applications* 2(4): 315-322 (2003); http://www.cs.unb.ca/ai2002/baseweb/BASeWEB2002_Paper6.pdf

[SollazzoHSFS:02] T. Sollazzo, Siegfried Handschuh, Steffen Staab, M. Frank, N. Stojanovic: "Semantic Web Service Architecture — Evolving Web Service Standards toward the Semantic Web"; *Proc. of the 15th International FLAIRS Conference*. Pensacola, Florida, May 16-18, 2002. AAAI Press.

[StaabEAD:01] Steffen Staab, Michael Erdmann, Alexander Mädche, Stefan Decker: "An extensible approach for Modeling Ontologies in RDF(S)"; *Knowledge Media in Healthcare: Opportunities and Challenges*. Rolf Grütter (ed.). Idea Group Publishing, Hershey USA / London, UK. December 2001.

[StaabMH:01] S. Staab, A. Maedche, S. Handschuh: "An Annotation Framework for the Semantic Web"; S. Ishizaki (ed.): *Proc. of The First International Workshop on MultiMedia Annotation*. January. 30 - 31, 2001. Tokyo, Japan.

[StaabS:04] Steffen Staab, Rudi Studer (eds.): Handbook on Ontologies; International Handbooks on Information Systems, Springer Verlag, 2004.

[StaabSS:02] Steffen Staab, Rudi Studer, Y. Sure: "Knowledge Processes and Meta Processes in Ontology-based Knowledge Management", C. Holsapple (ed.) *Handbook on Knowledge Management*. International Handbooks on Information Systems, Springer Verlag, 2002.

[StaabSSS:01] Steffen Staab, H.-P. Schnurr, Rudi Studer, Y. Sure: "Knowledge Processes and Ontologies"; *IEEE Intelligent Systems*. 16(1), January/February 2001. Special Issue on Knowledge Management.

[StaabSSV:02] Steffen Staab, Rudi Studer, Y. Sure, R. Volz: "SEAL - a SEmantic portAL with content management functionality"; *Gaining Insight from Research Information. CRIS 2002 - Proceedings of the 6th Conference on Current Research Information Systems*, August 29-31, 2002, Kassel, Germany.

[StojanovicSH:02] L. Stojanovic, N. Stojanovic, Siegfried Handschuh: Evolution of the Metadata; *Ontology-based Knowledge Management Systems*. German Workshop on Experience Management 2002.

[StuderDFS:03] Rudi Studer, Stefan Decker, Dieter Fensel, & Steffen Staab: "Situation and Prospective of Knowledge Engineering"; J.Cuena, Y.Demazeau, A. Garcia, J.Treur (eds.). Knowledge Engineering and Agent Technology. IOS Series on Frontiers in Artificial Intelligence and Applications, Volume 52, IOS Press, 2003.

[SureAS:02] York Sure, J. Angele, Steffen Staab: "OntoEdit: Guiding Ontology Development by Methodology and Inferencing"; *1st International Conference on Ontologies, Databases and Applications of Semantics for Large Scale Information Systems - ODBASE 2002*. October 29 - November 1, Irvine, California. LNCS, Springer, 2002.

[SureAS:03] Y. Sure, J. Angele, Steffen Staab: "OntoEdit: Multifaceted Inferencing for Ontology Engineering"; *Journal on Data Semantics*, LNCS 2800, Springer, 2003, pp. 128-152.

[SureEASSW:02] York Sure, Michael Erdmann, J. Angele, Steffen Staab, Rudi Studer, D.Wenke: "OntoEdit: Collaborative Ontology Development for the Semantic Web"; *Proceedings of the 1st International Semantic Web Conference - ISWC2002*, Springer, LNCS.

[SureS:02] York Sure and Rudi Studer, "On-To-Knowledge Methodology," in *On-To-Knowledge: Semantic Web enabled Knowledge Management*, J. Davies and D. Fensel and F. van Harmelen, Ed. 2002,

[SureSS:02] York Sure, Steffen Staab, Rudi Studer: "Methodology for Development and Employment of Ontology-based Knowledge Management Applications"; *Sigmod Record*, December 2002.

[SureSS:04] York Sure, Steffen Staab, Rudi Studer: "Methodology for Development and Employment of Ontology-based Knowledge Management Applications"; Steffen Staab & Rudi Studer (eds.) *Handbook on Ontologies*. Springer 2004.

[TangmunarunkitDK:03] Hongsuda Tangmunarunkit, Stefan Decker, Carl Kesselman: " Ontology-Based Resource Matching in the Grid - The Grid Meets the Semantic Web"; *International Semantic Web Conference 2003*: 706-721.

[VolzDC:03] Raphael Volz, Stefan Decker, Isabel F. Cruz: PSSS1 - Practical and Scalable Semantic Systems, *Proceedings of the First International Workshop on Practical and*

Scalable Semantic System; Sanibel Island, Florida, USA, October 20, 2003; http://km.aifb.uni-karlsruhe.de/ws/psss03/proceedings/

[VolzHSSS:04] R. Volz, Siegfried Handschuh, Steffen Staab, L. Stojanovic, N. Stojanovic: "Unveiling the Hidden Bride: Deep Annotation for Mapping and Migrating Legacy Data to the Semantic Web"; *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*; Vol.1, Elsevier, 2004 p.187-.206.

[Wiederhold:94] Gio Wiederhold: "An Algebra for Ontology Composition"; *Proceedings of 1994 Monterey Workshop on Formal Methods*, Sept 1994, U.S. Naval Postgraduate School, Monterey CA, pages 56-61.

[Wiederhold:00] Gio Wiederhold: "Precision in Processing Data from Heterogeneous Resources"; B.Lings and K.Jeffreys (eds.): *Advances in Databases*, Proc. 17th British National Conf. on Databases, Exeter, UK, July 2000, pages 1-18; http://www-db.stanford.edu/pub/gio/2001/BNCOD.doc

[Wiederhold:01] Gio Wiederhold: "The Need and Tools to Gain Precision in Electronic Commerce"; *Software Tech News*, Data and Analysis Center for Software (USAF, DoD), Rome, NY, Vol.4 No.4, Oct. 2001, pages 16-27.

[Wiederhold:02] Gio Wiederhold: "Obtaining Precision when Integrating Information"; J.Filipe, Sharp, B. and Miranda, P. (Eds.), *Enterprise Information Systems III*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002; http://www-db.stanford.edu/pub/gio/2001/ICEISprecision.doc.

[Wiederhold:03a] Gio Wiederhold: "The Impossibility of Global Consistency"; Position paper updated for the NLM/NSF workshop on A Research Challenge for Biological Data Management; Feb 2003, NLM, Bethesda MD; http://www-db.stanford.edu/pub/gio/2003/ConsistencyChallenge5Feb2003.htm.

[Wiederhold03b] Gio Wiederhold: "Increasing the Information Density in Digital Library Results"; Presentation Summary, Indo-US Workshop on Open Digital Libraries and Interoperability, June 2003; http://fox.cs.vt.edu/IndoUSdl/Wiederhold.pdf.

[WiederholdA:04] Gio Wiederhold and Fernando Arguello: References providing access to the XML data files and their DTD Schemas for the Movies Database; 2004, http://www-db.stanford.edu/pub/movies/dtd.html.

[Wiederhold:05] Gio Wiederhold: "What are Web Services Worth?", contribution to the Web business workshop if the 4th European Semantic Web Conference, May 2005; http://www-db.stanford.edu/pub/gio/inprogress.html#worth.